

# PATENT ABSTRACTS OF JAPAN

(11)Publication number : 2001-166787  
(43)Date of publication of application : 22.06.2001

(51)Int.Cl. G10L 13/06  
G06F 3/16  
G10L 13/08  
G10L 15/18

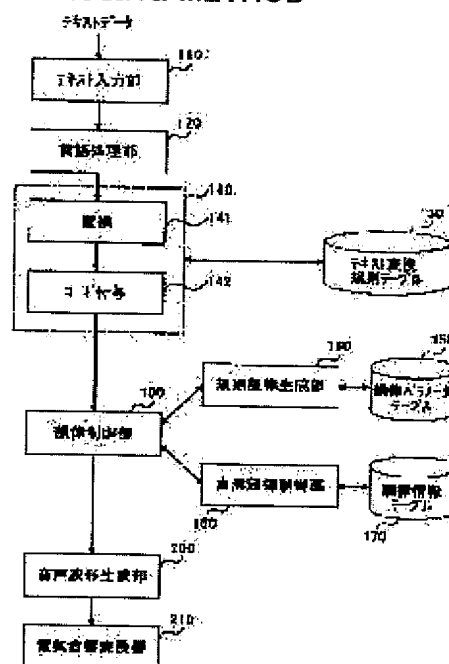
(21)Application number : 11-344959 (71)Applicant : MATSUSHITA ELECTRIC IND CO LTD  
(22)Date of filing : 03.12.1999 (72)Inventor : KATO YUMIKO  
YAMAGAMI KATSUYOSHI

## (54) VOICE SYNTHESIZER AND NATURAL LANGUAGE PROCESSING METHOD

### (57)Abstract:

**PROBLEM TO BE SOLVED:** To provide a voice synthesizer which can read a text of written words aloud with a rhythm close to that of spoken words by replacing the text with expression of the spoken words and read the text aloud naturally with the spoken words.

**SOLUTION:** A text conversion part 140 converts a part in an input text which corresponds to a pattern of written words of a text conversion rule registered in a text conversion rule table 130 into spoken words specified by the rule. A rhythm information control part 190 gives rhythm information which is specially prepared for the spoken words to the part converted into the spoken words and generates a rhythm other than the converted part according to the rule by using a parameter so set as to naturally connect with the converted part.



## LEGAL STATUS

[Date of request for examination]

[Date of sending the examiner's decision of rejection]

[Kind of final disposal of application other than the examiner's decision of rejection or application converted registration]

[Date of final disposal for application]

[Patent number]

[Date of registration]

[Number of appeal against examiner's decision of rejection]

[Date of requesting appeal against examiner's  
decision of rejection]

[Date of extinction of right]

(19) 日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11) 特許出願公開番号  
特開2001-166787  
(P2001-166787A)

(43) 公開日 平成13年6月22日 (2001.6.22)

(51) Int.Cl. <sup>7</sup>	識別記号	F I	テマコード* (参考)
G 1 0 L 13/06		G 0 6 F 3/16	3 3 0 K 5 D 0 1 5
G 0 6 F 3/16	3 3 0	G 1 0 L 5/04	F 5 D 0 4 5
G 1 0 L 13/08		3/00	H 9 A 0 0 1
15/18			5 3 7 Z

審査請求 未請求 請求項の数13 O L (全 14 頁)

(21) 出願番号 特願平11-344959

(22) 出願日 平成11年12月3日 (1999.12.3)

(71) 出願人 000005821

松下電器産業株式会社

大阪府門真市大字門真1006番地

(72) 発明者 加藤 弓子

大阪府門真市大字門真1006番地 松下電器  
産業株式会社内

(72) 発明者 山上 勝義

大阪府門真市大字門真1006番地 松下電器  
産業株式会社内

(74) 代理人 100105050

弁理士 鷲田 公一

Fターム(参考) 5D015 AA05

5D045 AA07 AA09

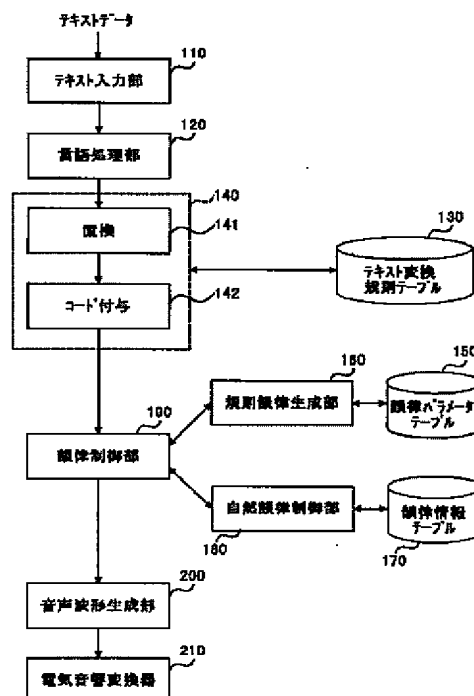
9A001 HH13 HH18 HH22

(54) 【発明の名称】 音声合成装置および自然言語処理方法

(57) 【要約】

【課題】 書き言葉のテキストを話し言葉の表現に置き換えて話し言葉らしい韻律で読み上げ、かつ、話し言葉韻律にあわせて規則韻律を調整することで、話し言葉として自然にテキストを読み上げることができる音声合成装置を提供すること。

【解決手段】 テキスト変換部140において、入力テキスト中で、テキスト変換規則テーブル130に登録されているテキスト変換規則の書き言葉のパターンに照合する部分を、規則に指定された話し言葉へと変換する。韻律情報制御部190においては、話し言葉へ変換された部分に対して、話し言葉のために特別に用意した韻律情報を付与する一方、変換部以外の韻律を変換部へ自然につながるよう設定されたパラメータを用いて規則に従って生成する。



## 【特許請求の範囲】

【請求項1】 入力テキストの表現を音声化に適した表現に変換する表現変換手段と、入力テキスト中の前記表現変換手段によって変換された部分について、話し言葉用の韻律情報を生成する自然韻律制御手段と、前記表現が変換された部分に接続する、表現が変換されない部分の韻律を調整するためのパラメータの値を適応的に決定することによって、前記表現が変換されない部分についての韻律情報を生成する規則韻律生成手段と、を有することを特徴とする、音声合成装置。

【請求項2】 前記規則韻律生成手段は、前記表現が変換された部分に接続する、表現が変換されない部分が、前記表現が変換された部分に自然につながるように、予め定められているパラメータ値を選択する、複数のパラメータ値の中から適切なものを選択する、あるいは、韻律目標値になるように所定の計算をする、のいずれかの手法を用いて、前記パラメータの値を決定することの特徴とする請求項1記載の音声合成装置。

【請求項3】 日本語の入力テキストを単語列に分割し、単語列の各単語に対して読み、品詞などの言語情報を付与する言語処理手段と、  
変換前テキストパターンと、変換後テキストと、前記変換後テキストを読み上げるための韻律情報を指定する韻律情報識別コード、および変換部以外のテキストを読み上げるための韻律情報を生成する際の1つ以上のパラメータを指定する規則韻律パラメータ識別コードを1組とするテキスト変換規則が登録されているテキスト変換規則テーブルと、

前記言語処理手段から出力される単語列と、前記テキスト変換規則テーブルを照合して、照合した部分をテキスト変換規則に従って書きかえるテキスト変換手段と、  
前記テキスト変換規則の規則韻律パラメータ識別コードで特定される1つ以上の規則韻律パラメータが登録されている規則韻律パラメータテーブルと、  
前記テキスト変換手段から出力される単語列に付与された言語情報から発音系列を生成し、前記テキスト変換規則の規則韻律パラメータコードで指定される1つ以上の規則韻律パラメータを規則韻律パラメータテーブルから取り出し、そのパラメータを用いて、あらかじめ定められた規則に従って前記発音系列に対応する韻律情報を生成する規則韻律生成手段と、

前記テキスト変換規則の韻律情報識別コードで特定される韻律情報が登録されている韻律情報テーブルと、  
前記テキスト変換手段で前記テキスト変換規則によって書き換えられた部分に対して、前記テキスト変換規則の韻律情報識別コードで指定される韻律情報を韻律情報テーブルから取り出す自然韻律制御手段と、  
前記規則韻律生成手段および前記自然韻律制御手段から生成された各韻律情報を統合することで、入力テキスト全体の韻律情報を生成する韻律制御手段と、

前記韻律情報制御手段で生成された韻律情報に従って音声波形を合成する音声波形生成手段と、を有することを特徴とする音声合成装置。

【請求項4】 日本語の入力テキストを単語列に分割し、単語列の各単語に対して読み、品詞などの言語情報を付与する言語処理手段と、  
変換前テキストパターンと、変換後テキストと、前記変換後テキストを読み上げるための韻律情報を指定する韻律情報識別コード、および変換部以外のテキストを読み上げるための韻律情報を生成する際の1つ以上の規則韻律制御パラメータを1組とするテキスト変換規則が登録されている、テキスト変換規則テーブルと、

前記言語処理手段から出力される単語列と、前記テキスト変換規則テーブルを照合して、照合した部分をテキスト変換規則に従って書きかえるテキスト変換手段と、  
前記テキスト変換手段から出力される単語列に付与された言語情報から発音系列を生成し、前記テキスト変換手段から出力される規則韻律制御パラメータを用いてあらかじめ定められた規則に従って前記発音系列に対応する韻律情報を生成する規則韻律生成手段と、  
前記テキスト変換規則の韻律情報識別コードで特定される韻律情報が登録されている韻律情報テーブルと、  
前記テキスト変換手段で前記テキスト変換規則によって書きかえられた部分に対して、前記テキスト変換規則の韻律情報識別コードで指定される韻律情報を韻律情報テーブルから取り出す自然韻律制御手段と、  
前記規則韻律生成手段と前記自然韻律制御手段から生成された韻律情報と統合することで、入力テキスト全体の韻律情報を生成する韻律制御手段と、  
前記韻律情報制御手段で生成された韻律情報に従って音声波形を合成する音声波形生成手段と、を有することを特徴とする音声合成装置。

【請求項5】 前記規則韻律パラメータテーブルあるいは前記テキスト変換規則テーブルは、前記規則韻律生成手段で韻律情報を生成する、前記テキスト変換がなされなかった部分のテキストの長さに応じて使い分けがなされる規則韻律制御パラメータを、複数保持していることを特徴とする請求項3または請求項4記載の音声合成装置。

【請求項6】 前記規則韻律生成手段で韻律情報を生成する変換部以外のテキストの長さはモーラ数によって判断するものとする請求項5記載の音声合成装置。

【請求項7】 前記規則韻律生成手段で韻律情報を生成する、前記テキスト変換がなされなかった部分のテキストの長さは、アクセント句数によって判断することを特徴とする請求項5記載の音声合成装置。

【請求項8】 日本語の入力テキストを単語列に分割し、単語列の各単語に対して読み、品詞などの言語情報を付与する言語処理手段と、  
変換前テキストパターンと、変換後テキストと、前記変換後テキストを読み上げ

10

20

30

40

50

るための韻律情報を指定する韻律情報識別コード、および交換部と交換部以外との接続部分での韻律情報である 1 つ以上の規則韻律目標値を 1 組とするテキスト交換規則が登録されているテキスト交換規則テーブルと、前記言語処理手段から出力される単語列と、前記テキスト交換規則テーブルを照合して、照合した部分をテキスト交換規則に従って書きかえるテキスト交換手段と、前記テキスト交換手段から出力される単語列に付与された言語情報から発音系列を生成し、その発音系列のうちテキスト交換部との接続部分が前記テキスト交換手段から出力される規則韻律目標値となるよう規則韻律パラメータを決定する規則韻律パラメータ計算手段と、前記発音系列に従ってあらかじめ定められた規則に従って韻律情報を生成する規則韻律生成手段と、前記テキスト交換規則の韻律情報識別コードで特定される韻律情報が登録されている韻律情報テーブルと、前記テキスト交換手段で前記テキスト交換規則によって書きかえられた部分に対して、前記テキスト交換規則の韻律情報識別コードで指定される韻律情報を韻律情報テーブルから取り出す自然韻律制御手段と、前記規則韻律生成手段と前記自然韻律制御手段から生成された韻律情報と統合することで、入力テキスト全体の韻律情報を生成する韻律制御手段と、前記韻律情報制御手段で生成された韻律情報に従って音声波形を合成する音声波形生成手段と、を有することを特徴とする音声合成装置。

【請求項 9】 前記テキスト交換規則テーブルに登録されているテキスト交換規則において、交換前テキストパターンを書き言葉の表現パターンとし、交換後テキストを話し言葉の表現とし、韻律情報識別コードを話し言葉の韻律情報を指定するコードとすることを特徴とする、請求項 3、請求項 5 または請求項 8 のいずれかに記載の音声合成装置。

【請求項 10】 前記韻律情報テーブルは、話し言葉特有の表現を読み上げるための韻律情報を保持することを特徴とする、請求項 3、請求項 4 または請求項 8 のいずれかに記載の音声合成装置。

【請求項 11】 前記韻律情報は、基本周波数、強さ、話速、音韻時間長の情報のうち 1 つ以上を含むものであることを特徴とする、請求項 3、請求項 4 または請求項 8 のいずれかに記載の音声合成装置。

【請求項 12】 入力テキストを形態素解析して品詞などの言語情報を付与した単語列に分解するステップと、前記単語列の中の、テキスト交換規則テーブルに登録されている単語列パターンに合致するものを話し言葉の単語列に置き換えるステップと、その置き換えられた単語列に対して、自然韻律情報を生成するステップと、前記置き換えられた単語列に連なる、置換がなされない部分の韻律を調整するためのパラメータの値を適応的に決定することによって、前記置換

がなされない部分についての規則韻律情報を生成するステップと、

前記自然韻律情報および規則韻律情報に基づいて、前記入力テキストのイントネーションを生成することの特徴とする自然言語処理方法。

【請求項 13】 前記パラメータの適応的決定は、単語列の置き換えがなされなかった部分が、前記単語列の置き換えがなされた部分に自然につながるように、予め定められているパラメータ値を選択する、複数のパラメータ値の中から適切なものを選択する、あるいは、入力されたテキストの言語情報に基づいて所定の計算をする、のいずれかの手法を用いて行なわれることを特徴とする請求項 12 記載の自然言語処理方法。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】本発明は、入力された任意のテキストの読み上げ音声合成するテキスト音声合成装置、および、テキスト音声合成装置において書き言葉を話し言葉に変換して読み上げるための自然言語処理方法に関する。

【0002】

【従来の技術】従来の音声合成装置は、日本語テキストを単純に分割し、発音記号列に変換し韻律情報を生成してその韻律にしたがって波形を生成していた。しかし、テキスト情報は読まれることを前提とした書き言葉で記述されている場合がほとんどであり、発音しにくい、あるいは、聞き取りにくい単語、複合語、フレーズを含んでいることがあり、そのまま読み上げた場合には、合成音声の理解性を損なっていた。

【0003】この点に着目し、入力テキストの難意語や紛らわしい同音異義語を平易な単語に置き換える機能を有したテキスト音声合成装置（特開平 3-35296）や、聞いてわかりやすく自然な音声にするため、テキストの表現を話し言葉に適した形式に変換して音声合成するという音声合成技術が提案されている。

【0004】

【発明が解決しようとする課題】しかしながら、話し言葉の文末や区切れ、意味の切れ目等では、自然音声では基本周波数、音声強度、話速や音韻時間長が複雑に調節されており、規則に基づく音声合成方式では自然な音声を生成することは、なかなか困難である。

【0005】すなわち、話し言葉特有の表現に見られる韻律パターンは複雑であり規則音声合成で再現することは困難である。そのため書き言葉のテキストを話し言葉に変換して音声を合成しても、自然性、理解性は改善されないばかりか、低下する場合も想定される。

【0006】実際、ラジオなどのニュースを読み上げるアナウンサーは、書き言葉から話し言葉へと修正された原稿を、話し言葉特有の表現をそれぞれの表現に適したイントネーションで読み上げることで、意味のまとま

り、区切れを明確にし、内容をわかりやすく正確に伝える。したがって、仮に、書き言葉を話し言葉に変換し、その部分だけを自然な音声（例えば、アナウンサー特有のイントネーション）で読み上げても、その部分に接続する（連続する）部分が機械的な音声であって自然に滑らかに続いていないならば、かえって、読み上げの調子が途中で唐突に変化して、かえって、聴取者に違和感を与えることになる。

【0007】本発明は、このような不都合を解消するためになされたものであり、自然で了解性の高い音声でテキストの読み上げを行うことができる音声合成装置および自然言語処理方法を提供することを目的とする。

【0008】

【課題を解決するための手段】本発明では、書き言葉を話し言葉に変換して自然韻律（例えば、話し手であるアナウンサー特有の自然なイントネーション）で読み上げるのみならず、その変換部分に接続する部分についても、読み上げを調節するパラメータを調節して、変換部分にできるだけ自然につながるように規則合成する。

【0009】変換部分以外の部分についてパラメータを調整する方法としては、変換部分の言葉に連続させるのにふさわしいパラメータをあらかじめ一つだけ決めておく方法や、文章の長さに応じて、複数の候補の中から一つを選択する方法や、最もふさわしい韻律を目標値として与え、その目標値になるようにパラメータをリアルタイムで計算する方法などがある。

【0010】

【発明の実施の形態】本発明の音声合成装置の一つの態様では、書き言葉の表現を話し言葉特有の表現に変換し、その変換部に対応する話し言葉特有の韻律情報コードを付与し、またその話し言葉特有の韻律に接続するための韻律制御パラメータコードを付与し、それらのコードより変換部には話し言葉特有の自然な韻律情報を適用し、変換部以外には変換部の話し言葉特有の韻律に自然に接続する韻律情報を生成するためのパラメータをコードに対応して適用し、規則に従って韻律情報を生成して両者を統合し、その韻律に基づいて音声を合成する。

【0011】また、本発明の音声合成装置の他の態様では、書き言葉の表現を話し言葉特有の表現に変換し、その変換部に対応する話し言葉特有の韻律情報コードを付与し、またその話し言葉特有の韻律に接続するための韻律制御パラメータを、例えば、読み上げしようとする文書の長さに応じて付与し、変換部には話し言葉特有の韻律情報コードに対応する話し言葉特有の自然な韻律情報を適用し、変換部以外には話し言葉特有の韻律に接続するための韻律制御パラメータを用いて、変換部の話し言葉特有の韻律に自然に接続する韻律情報を規則に従って生成して両者を統合し、その韻律に基づいて音声を合成する。

【0012】また、本発明の音声合成装置の他の態様で

は、書き言葉の表現を話し言葉特有の表現に変換し、その変換部に対応する話し言葉特有の韻律情報コードを付与し、またその話し言葉特有の韻律に接続する部分の韻律情報を変換部以外の韻律を生成する際の目標値として付与し、変換部には話し言葉特有の韻律情報コードに対応する話し言葉特有の自然な韻律情報を適用し、変換部以外には変換部との接続部で韻律情報の目標値を実現するようなパラメータを設定しそのパラメータを用いて韻律情報を規則に従って生成して両者を統合し、その韻律に基づいて音声を合成する。

【0013】これらにより、書き言葉で書かれた入力テキストを話し言葉の表現に置き換えて話し言葉らしい韻律で読み上げると共に、話し言葉特有でない表現の部分との接続を自然にすることが可能であり、全体としてまとまりのある自然な合成音声を提供することができる。

【0014】以下、本発明の実施の形態について、図面を参照して具体的に説明する。

【0015】（実施の形態1）図1は、本発明の実施の形態1における音声合成装置の構成を示すブロック図である。

【0016】本実施の形態では、テキスト変換規則テーブルに自然音声韻律パターンコードおよび規則合成用韻律用パラメータコードを記憶させておき、テキスト変換の際に、変換部分の自然韻律パタンのみならず、変換部分以外の、変換部分に連なる部分のパラメータもコードにより指定してしまうものである。

【0017】図1に示すように、本実施の形態の音声合成装置は、テキスト入力部110と、入力された日本語テキストを解析する言語処理部120と、書き言葉と話し言葉の変換規則および変換部の韻律情報コードおよび規則合成のパラメータコードを記録したテキスト変換規則テーブル130と、テキスト変換規則テーブル130を参照して、書き言葉を話し言葉に変換するテキスト変換部140（機能ブロックとして置換手段141と、コード付与手段142をもつ）と、を有する。

【0018】さらに、この音声合成装置は、規則合成のパラメータコードに対応する基本周波数制御パラメータPt、音声強度制御パラメータPw、話速制御パラメータSr、のセットを記録している韻律パラメータテーブル150と、韻律パラメータテーブル150を参照して、あらかじめ定められた規則により変換部以外の韻律情報を生成する規則韻律生成部160と、自然韻律パタンコードと基本周波数パタン、音声強度パタン、音韻時間長パタンを記録した自然韻律情報テーブル170と、自然韻律情報テーブル170を参照して変換部の韻律情報を生成する自然韻律制御部180と、規則韻律生成部160から出力される変換部以外の韻律情報と自然韻律制御部180から出力される変換部の韻律情報とを統合する韻律制御部190と、音波形成部200と、電気音響機器210と、を有する。

【0019】なお、音声波形生成部200は、韻律制御部190から出力される韻律情報と発音記号列とに基づき音声波形を生成する。

【0020】次に、音声合成装置の動作を説明する。

【0021】テキスト入力部110は、処理すべきテキストデータを受け付け、言語処理部120は、入力されたテキストデータを形態素解析して、品詞などの言語情報を付与した単語列に分割する。

【0022】テキスト変換部140は、言語処理部120の出力である単語列を走査して、テキスト変換規則テーブル130の変換前の単語列パターンに照合する部分を、変換後の単語列に置き換えて、その部分の単語列には自然韻律情報識別コードを付与し、一方、変換を行わない部分に対しては規則合成用韻律パラメータコードを付与する。テキスト変換規則テーブル130には、入力テキストの単語列と照合すべきテキスト変換規則が複数登録されている。テキスト変換規則テーブル130の内容の一例を図2に示す。

【0023】例えば、「…の開設を検討するという。」というテキストが入力された場合を想定する。この場合、「という。」は「ということです。」に変換され、韻律情報識別コード「0001」が付与される。また、この変換された部分に連なる「開設を検討する」という部分に対しては、規則合成用韻律パラメータコード「0001」が付与される。この規則合成用韻律パラメータコードは、変換された部分に最も自然につながる合成音声出力するべく、規則韻律のパラメータの組を指定するものである。このパラメータコードは、表現が書き言葉から話し言葉に変換される部分に対応させて、予め選択されているものである。

【0024】図3に、テキスト変換部140による変換の前後のテキストデータの例を示す。図中、参照符号401で示される部分が変換前単語列であり、参照符号402で示される部分が、変換後の、変換部分に接続する部分の単語列であり、参照符号403で示される部分が、変換部分の単語列である。

【0025】図3の下側に示されるような単語列が、テキスト変換部140から出力され、韻律制御部190に入力される。韻律制御部190は、入力された単語列のうち、変換部に相当する部分を自然韻律制御部180に与え、変換部以外の部分の単語列を規則韻律生成部160に与える。そして、規則韻律生成部160から出力される変換部以外の部分に対応する韻律情報と、自然韻律制御部180から出力される変換部の韻律情報を統合する。

【0026】規則韻律生成部160は、上述のテキスト変換部140から出力される単語列に付与された言語情報と規則合成用韻律パラメータコードに基づき、韻律パラメータテーブル150を参照して、基本周波数制御パラメータPt0001、音声強度制御パラメータPw000

1、話速制御パラメータSr0001、を選択し、テキスト変換部140で変換されなかった部分の韻律情報をあらかじめ定められた規則にしたがって生成する。なお、規則韻律生成部160は、例えば、藤崎モデルなどの規則化された韻律制御モデルによってピッチなどの韻律情報を生成する。

【0027】一方、自然韻律制御部180は、変換された部分に付与された自然韻律情報識別コードで指定される韻律情報を、韻律情報テーブル170から取り出し、テキスト変換部140で変換された部分の韻律情報を生成する。自然韻律情報テーブル170には、話し言葉の定型部分の韻律情報を登録しておく。この韻律情報は、例えば、実際に話し言葉を読み上げた音声を録音し、音声信号を分析して得られた韻律情報である。

【0028】上述したように、規則韻律生成部160から出力される変換部以外の部分に対応する韻律情報と、自然韻律制御部180から出力される変換部の韻律情報は、韻律制御部190で統合され、音声波形生成部200に送られる。

【0029】音声波形生成部200では、韻律制御部190で生成された韻律情報に従って音声波形を合成し、続いて、電気音響変換器210は、音声波形を音声に変換して出力する。

【0030】以上の動作をまとめると、図4に示すようになる。すなわち、書き言葉を話し言葉に変換する（ステップ220）。このとき、自然韻律情報識別コードと規則合成用のパラメータコードが付与される。

【0031】次に、規則合成のパラメータコードに対応する一組のパラメータ（Pt、Pw、Sr）を用いて、変換部分以外の部分について、規則韻律情報を生成する（ステップ221）と共に、基本周波数パターン、音声強度パターン、音韻時間長パターンを用いて、変換部分について、自然韻律情報を生成する（ステップ222）。

【0032】以上のように、本実施の形態の音声合成装置によれば入力テキストの中で話し言葉に変換すべき部分を、テキスト変換規則によって変換し、変換された部分について実際の音声信号の分析により得られる自然性の高い韻律情報を用いて音声合成を行い、変換部以外の韻律を自然韻律に合わせたパラメータで生成することにより、話し言葉特有の韻律を再現することができ、かつ、規則合成と自然韻律が自然につながる合成音声を提供することができる。

【0033】（実施の形態2）図5は、実施の形態2における音声合成装置の構成を示すブロック図である。

【0034】本実施の形態の音声合成装置の構成は、前掲の実施の形態のものとはほぼ同じである。但し、図3においては、テキスト変換規則テーブル131、テキスト変換部143、規則韻律生成部161を使用すると共に、図1の韻律パラメータテーブル150を使用しない

10

20

30

40

50

点において、異なる。

【0035】本実施の形態では、テキストの長さに応じて、規則韻律用のパラメータの組を適応的に変化させる。テキストの長さを規定する単位としては、ここでは、モーラ数を用いる。例えば、「博覧会の開設を検討するという」という短いテキストを音読する場合と、「もし、各国の首脳の合意が締結された場合には、博覧会の開設を検討するという」という長いテキストを音読する場合とでは、韻律が異なるのが通常である。例えば、長いテキストの場合、文頭から文末にいくにしたがって、話速は増大し、一番最後にやや速度を落として終了する、というような速度の変化がみられるし、文章のつながり目等で、声の調子が変わることもある。

【0036】これらのことを考慮し、テキストの文章の長さに応じて、規則韻律のパラメータを適応的に変化させ（具体的には、複数の候補の中から最も適したものを選択する）、実施の形態1よりも、さらに自然な音声合成を行なうものである。

【0037】以下、本実施の形態の音声合成装置の動作を説明する。

【0038】テキスト入力部110は、処理すべきテキストデータを受け付け、言語処理部120は、入力されたテキストデータを形態素解析して、品詞などの言語情報を付与した単語列に分割する。

【0039】テキスト変換部143は、言語処理部120の出力である単語列を走査して、テキスト変換規則テーブル131の変換前の単語列パターンに照合する部分を、変換後の単語列に置き換えて、その部分の単語列に自然韻律情報識別コードを付与すると共に、変換を行わない部分に対してそのモーラ数に応じて規則合成用韻律パラメータである基本周波数制御パラメータ $P_t$ 、音声強度制御パラメータ $P_w$ 、話速制御パラメータ $S_r$ 、リズムパラメータ $R_t$ 、を付与する。

【0040】図6に示すように「開設を検討するという。」というテキストが入力された場合、「という。」は「ということです。」に変換され、自然韻律情報識別コード「0001」が付与される。一方、「開設を検討する」はモーラ数が11であり、これに応じて規則合成用韻律パラメータ $P_{t131}$ 、 $P_{t132}$ 、 $\dots$ 、 $R_{t131}$ 、 $R_{t132}$ が付与される。規則韻律生成部161は、テキスト変換部143から出力される単語列に付与された言語情報と規則合成用韻律パラメータに基づいて、テキスト変換部143で変換されなかった部分の韻律情報をあらかじめ定められた規則にしたがって生成し、自然韻律制御部180は、変換された部分に付与された自然韻律情報識別コードで指定される韻律情報を、自然韻律情報テーブル170から取り出してテキスト変換部143で変換された部分の韻律情報を生成する。

【0041】韻律制御部190は、規則韻律生成部161から出力される変換部以外に対応する韻律情報と、自

然韻律制御部180から出力される変換部の韻律情報を統合する。音声波形生成部200は、韻律制御部190で生成された韻律情報に従って音声波形を合成し、電気音響変換器は音声波形を音声に変換して出力する。

【0042】以上の動作をまとめると、図7に示すようになる。すなわち、まず、単語列の変換と、自然韻律コードの付与および、モーラ数に応じた規則合成の韻律パラメータの付与を行なう（ステップ230）。

【0043】次に、規則合成の韻律パラメータに基づいて、規則韻律情報を生成すると共に（ステップ231）、自然韻律情報識別コードで指定される自然韻律情報を生成する（ステップ232）。そして、各情報を統合し（ステップ233）、音声の合成を行なう（ステップ234）。

【0044】以上のように、本実施の形態の音声合成装置によれば入力テキストの中で話し言葉に変換すべき部分を、テキスト変換規則によって変換し、変換された部分について実際の音声信号の分析により得られる自然性の高い韻律情報を用いて音声合成を行い、変換部以外の韻律を、入力テキストの長さ等に応じて適宜選択して、自然韻律に合わせたパラメータで生成することにより、話し言葉特有の韻律を再現することができ、規則合成と自然韻律が、より自然につながる合成音声を提供することができる。

【0045】（実施の形態3）図8は、本発明における音声合成装置の構成を示すブロック図である。

【0046】全体構成は、前掲の実施の形態とほぼ同じであるが、本実施の形態では、8においては、テキスト変換規則テーブル132と、テキスト変換部144と、規則韻律生成部161（規則韻律用パラメータ計算部360および韻律情報生成部460を含む）と、を使用する点で異なる。

【0047】本実施の形態の特徴は、変換を行なった部分に、変換を行なわない部分が最も自然につながるようにするべく、変換を行なわない部分の接続部分における最も好ましい韻律情報を目標値として付与し、この目標値の韻律情報を生成するためのパラメータ値をリアルタイムで計算することである。

【0048】以下、本実施の形態の音声合成装置の動作を説明する。

【0049】テキスト入力部110は、処理すべきテキストデータを受け付け、言語処理部120は、入力されたテキストデータを形態素解析して、品詞などの言語情報を付与した単語列に分割する。

【0050】テキスト変換部140は、言語処理部120の出力である単語列を走査して、テキスト変換規則テーブル132の変換前の単語列パターンに照合する部分を、変換後の単語列に置き換えて、その部分の単語列に自然韻律情報識別コードを付与し、変換を行わない部分に対して変換部との接続部分の韻律情報である規則合成



韻律目標値を付与する。

【0051】図9に示すように、「開設を検討するという。」というテキストが入力された場合、「という。」は「ということです。」に変換され、韻律情報識別コード「0001」が付与される。「開設を検討する」に対しては規則合成韻律目標値として図9のような、基本周波数、パワー、話速の情報が付与される。

【0052】規則韻律生成部161では、まず規則韻律用パラメータ計算部360で、テキスト変換部140cから出力される単語列に付与された言語情報から変換部10

以外のモーラ数やアクセント句数、呼吸段落の構成等を判断し、変換部との接続部の韻律がテキスト変換部144から出力される規則合成韻律目標値になるよう規則韻律用のパラメータを設定する。

【0053】韻律情報生成部460は、規則韻律用パラメータ計算部360から出力されるパラメータと、単語列に付与された言語情報から変換部以外の韻律情報をあらかじめ定められた規則にしたがって生成する。

【0054】自然韻律制御部180は、変換された部分に付与された自然韻律情報識別コードで指定される韻律20

情報を韻律情報テーブル170から取り出してテキスト変換部144で変換された部分の韻律情報を生成する。

【0055】韻律制御部190は、韻律情報生成部460から出力される変換部以外に対応する韻律情報と、自然韻律制御部180から出力される変換部の韻律情報を統合する。

【0056】音声波形生成部200は、韻律制御部190で生成された韻律情報に従って音声波形を合成する。

電気音響変換器210は、音声波形を音声に変換して出力する。

【0057】以上の動作をまとめると、図10のようになる。

【0058】すなわち、単語列の変換、自然韻律情報識別コードの付与および規則合成韻律目標値を付与する（ステップ240）。次に、変換部との接続部分の周波数、話速、パワー等が目標値と一致するように、パラメータを計算、出力し、規則韻律情報と自然韻律情報を生成する（ステップ241）。そして、各情報を統合し（ステップ242）、音声合成を行なう（ステップ243）。

【0059】以上のように、本実施の形態の音声合成装置によれば入力テキストの中で話し言葉に変換すべき部分を、テキスト変換規則によって変換し、変換された部分について実際の音声信号の分析により得られる自然性の高い韻律情報を用いて音声合成を行い、変換部と変換部以外の接続部の韻律目標値にあわせて規則韻律を生成することにより、話し言葉特有の韻律を再現することができ、規則合成と自然韻律が、最も自然につながる合成音声を提供することができる。

【0060】以上説明したように、本発明は、自然韻律

情報と、適応的なパラメータ制御で生成される規則韻律情報（適応型規則韻律情報）とを組み合わせて、自然な音声合成を行なうものである。これにより、図11に示すように、パワーやピッチが適切に調節され、表現が変換された部分（B）と、それに接続する部分（A）とを、自然につなげることができる。

【0061】なお、上述した構成の音声合成装置は、例えば、図12に示すようなコンピュータシステム上に構築されるものである。このコンピュータシステムは、本体部601と、キーボード602と、ディスプレイ603と、入力装置（マウス）604と、を含む、音声出力が可能なシステムである。テキスト変換規則テーブルや韻律情報テーブル等は、本体部601にセットされるCD-ROM607内、本体部601が内蔵するディスク（メモリ）606内、あるいは、回線608で接続された他のシステムのディスク605内に格納される。

【0062】以上、本発明について3つの実施の形態を用いて説明したが、本発明は、これに限定されるものではなく、種々、変形、応用が可能である。例えば、規則韻律生成は、藤崎モデルによるとしたが、それ以外のモデルに基づく規則であってもよい。

【0063】また、実施の形態の説明では、規則韻律生成のパラメータは、基本周波数パラメータ、音声強度パラメータ、話速パラメータとしたが、これにリズムパラメータを加えた4種類のパラメータのうち1つ以上を含むものであればよい。同じく、規則韻律生成のパラメータは基本周波数、音声強度、話速についてそれぞれ1つないしは2つのパラメータとして説明したが、それ以上の数のパラメータを指定してもよい。

【0064】また、実施の形態1では、テキスト変換を行わない部分の規則韻律用パラメータは合成音声を生じようとするテキストの長さとは無関係に設定したが、実施の形態2のように合成音性を生じようとするテキストの長さに応じて規則韻律用パラメータを設定してもよい。

【0065】また、実施の形態2において、テキスト変換を行わない部分の規則韻律用パラメータは合成音声を生じようとするテキストのモーラ数に応じて設定するとしたが、アクセント句数、文字数、単語数、文節数等他の単位を用いてもよく、さらに、これらの単位を複数組み合わせて設定してもよい。

【0066】また、実施の形態3において、規則合成韻律目標値は接続部の基本周波数、接続音韻最大パワー、接続部話速として説明したが、これ以外の韻律情報を含めてもよく、韻律情報を1つ以上含むものであればよい。

【0067】

【発明の効果】以上説明したように本発明によれば、書き言葉で書かれた入力テキストを話し言葉の表現に置き換えて話し言葉らしい韻律で読み上げ、かつ話し言葉特

有でない表現の部分との接続を自然にすることが可能であり、全体としてまとまりのある自然な合成音声を提供することができる。

【図面の簡単な説明】

【図1】本発明の実施の形態1にかかる音声合成装置のブロック図

【図2】本発明の実施の形態1のテキスト変換規則の内容の例を示す図

【図3】表現変換の前後の、単語列および付与される情報の例を示す図

【図4】実施の形態1にかかる音声合成装置の基本的な動作手順を示すフロー図

【図5】本発明の実施の形態2にかかる音声合成装置のブロック図

【図6】本発明の実施の形態2のテキスト変換規則の内容の例を示す図

【図7】実施の形態2にかかる音声合成装置の基本的な動作手順を示すフロー図

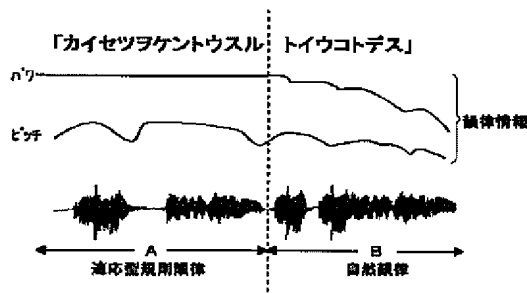
【図8】本発明の実施の形態3にかかる音声合成装置のブロック図

【図9】実施の形態3にかかるテキスト変換規則の内容\*

【図2】

入力テキスト	変換後テキスト	自然音声韻律パラメータ	規則合成用韻律パラメータ
「した。」	「しました。」	0000	0008
「という。」	「ということです。」	0001	0001

【図11】



\*の例を示す図

【図10】実施の形態3にかかる音声合成装置の基本的な動作手順を示すフロー図

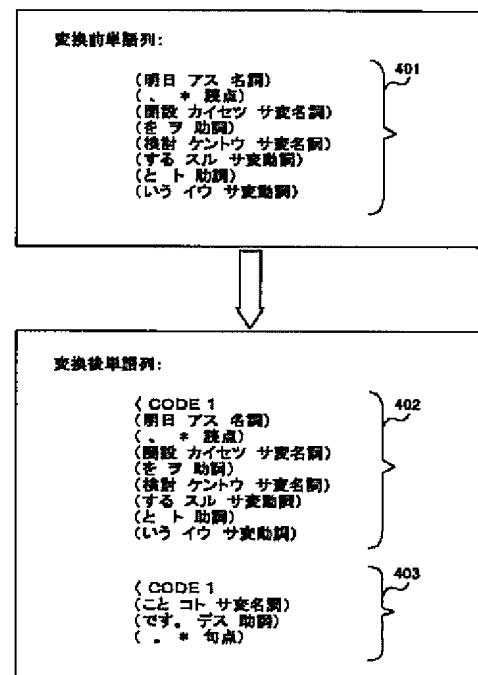
【図11】本発明の音声合成装置を用いた韻律生成の例を示す図

【図12】本発明を実施するためのコンピュータシステムの一例を示す図

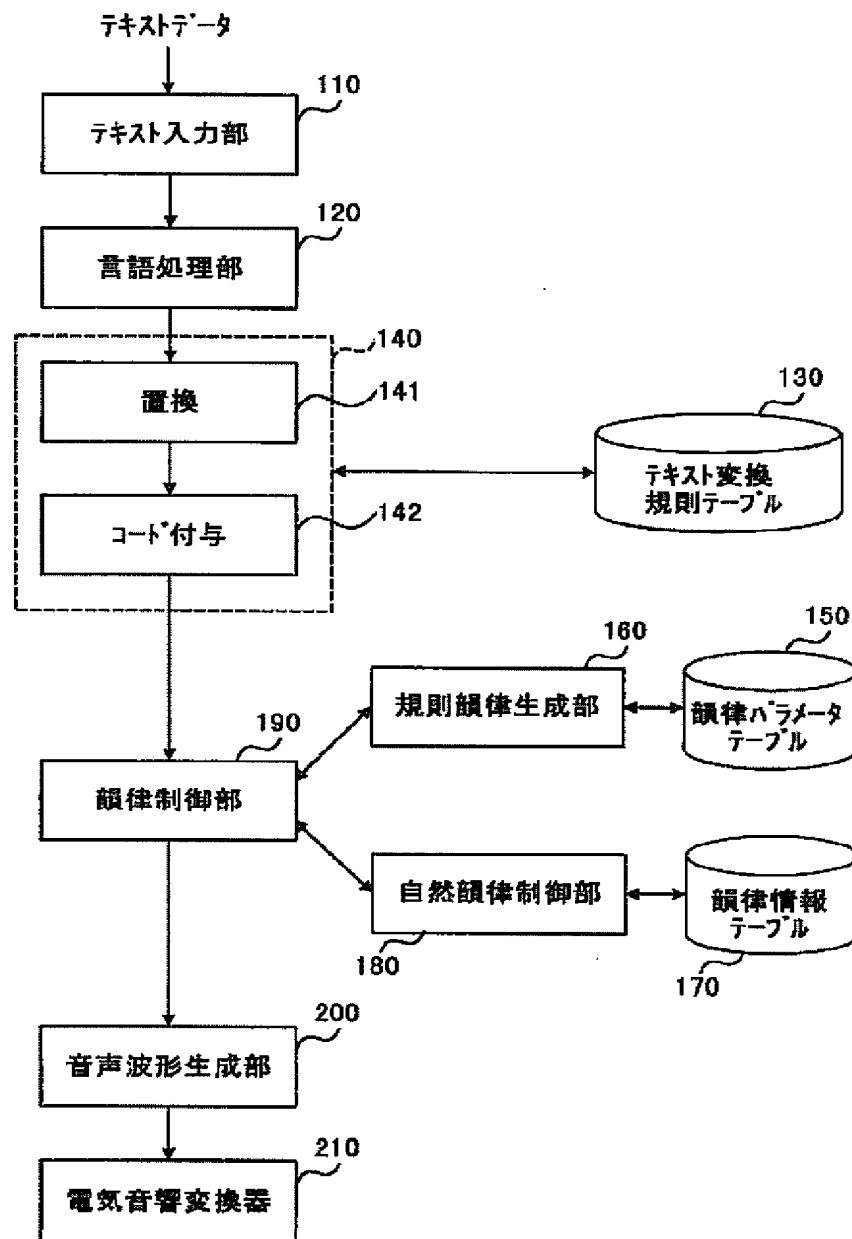
【符号の説明】

- 110 テキスト入力部
- 120 言語処理部
- 130 テキスト変換規則テーブル
- 140 テキスト変換部
- 141 表現置換部
- 142 コード付与部
- 150 韻律パラメータテーブル
- 160 規則韻律生成部
- 170 韻律情報テーブル
- 180 自然韻律制御部
- 190 韻律制御部
- 200 音声波形生成部
- 210 電気音響変換器

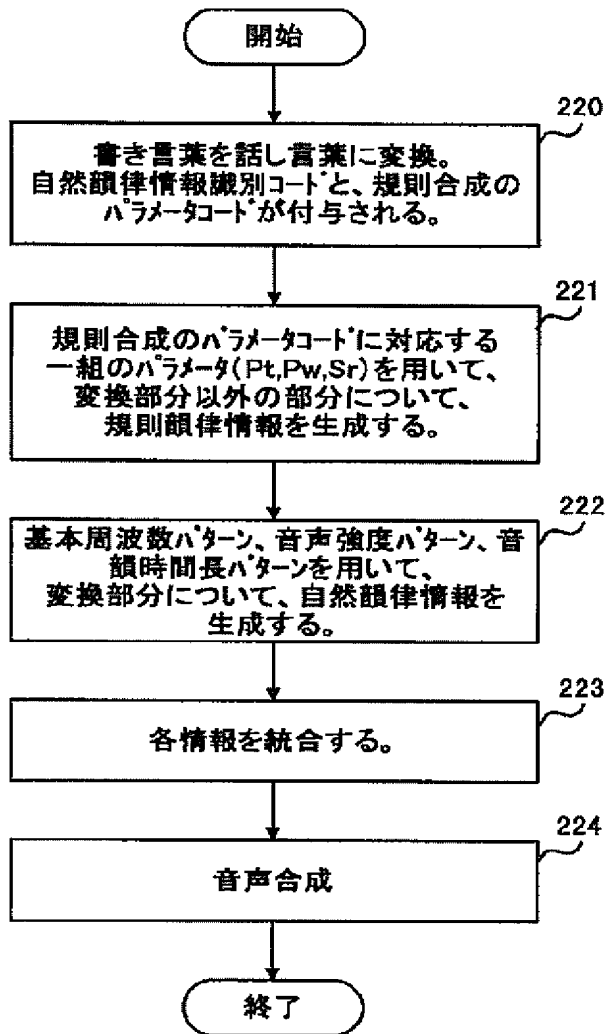
【図3】



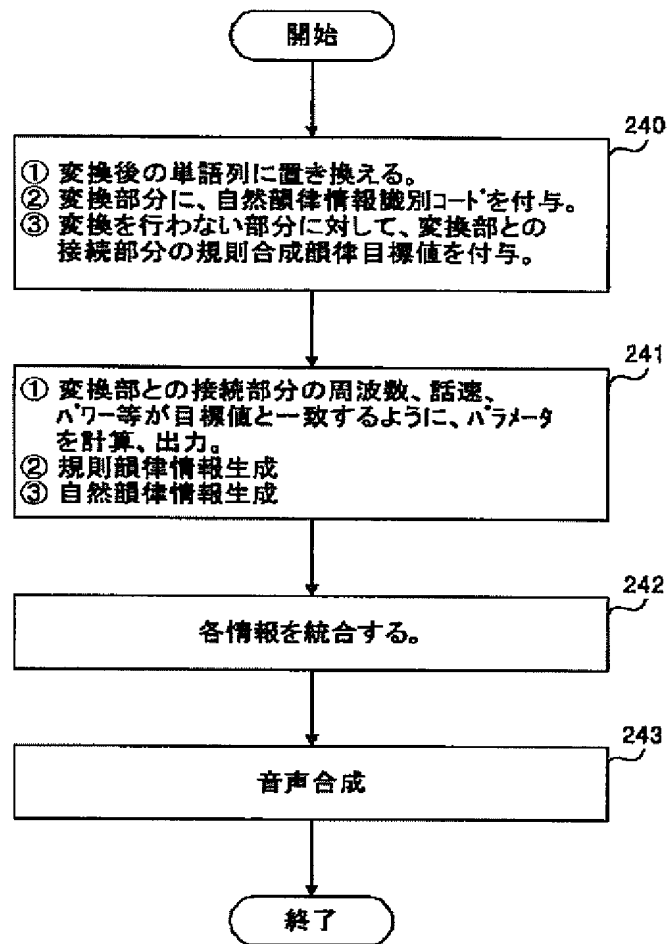
【図1】



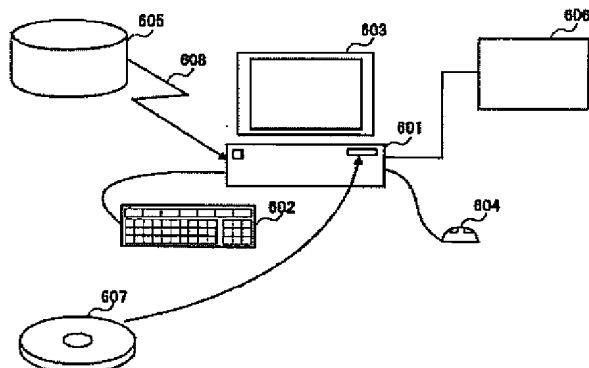
【図4】



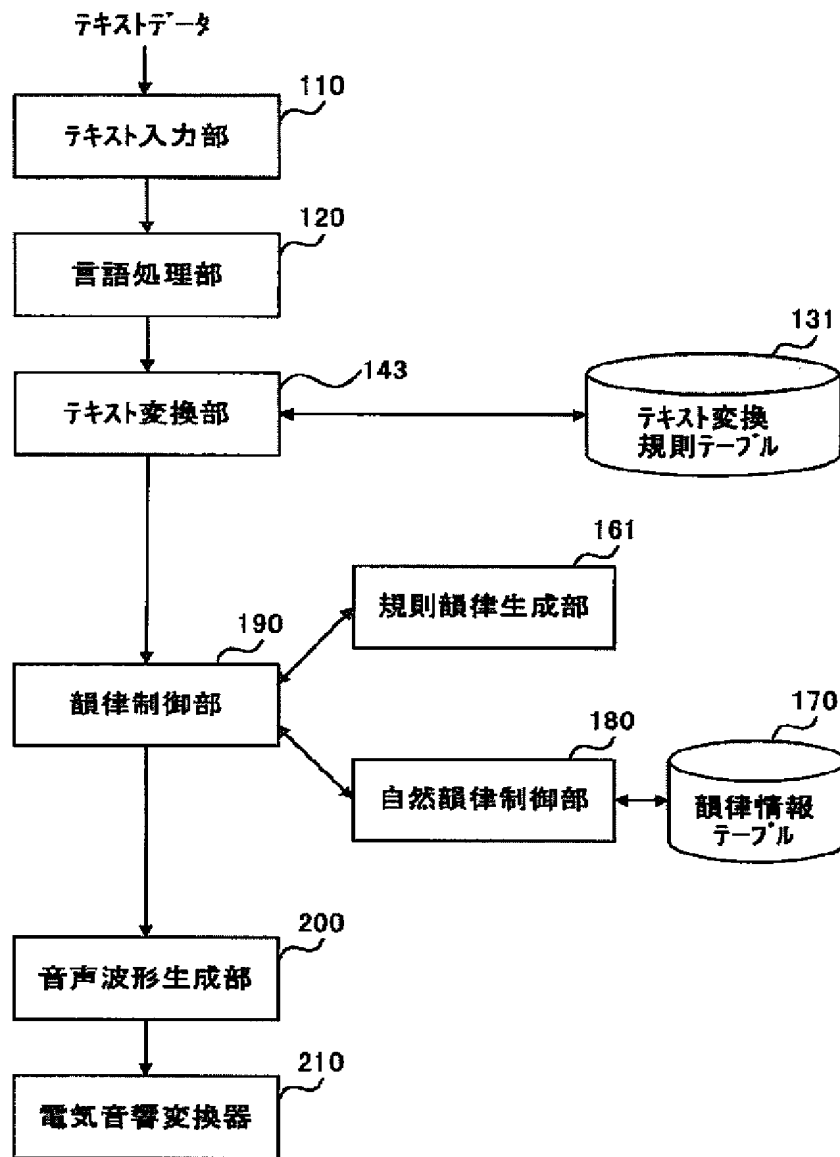
【図10】



【図12】



【図5】



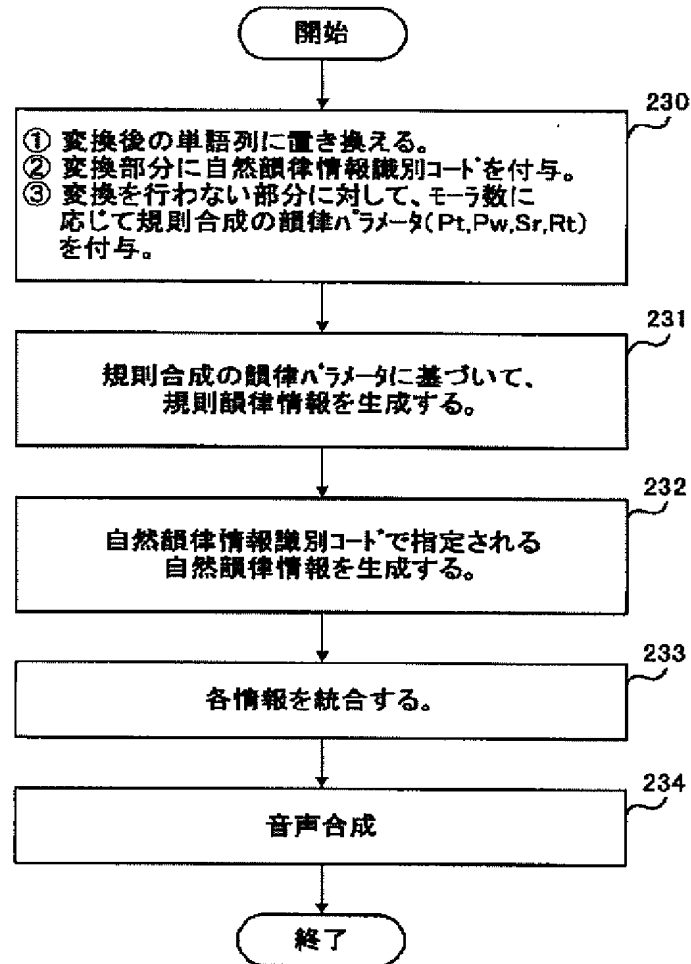
【図6】

入力テキスト	変換後テキスト	自然音声韻律 パターンコード	規則合成制御用韻律パラメータ (基本周波数パラメータ・パワーパラメータ・ 話速パラメータ・リズムパラメータ)		
			変換部以外の モーラ数>30	変換部以外の モーラ数>15	変換部以外の モーラ数≤15
「した。」	「しました。」	0000	Pt011,Pt012 Pw011,Pw012 Sr011,Sr012 Rt011,Rt012	Pt021,Pt022 Pw021,Pw022 Sr021,Sr022 Rt021,Rt022	Pt031,Pt032 Pw031,Pw032 Sr031,Sr032 Rt031,Rt032
「という。」	「ということです。」	0001	Pt111,Pt112 Pw111,Pw112 Sr111,Sr112 Rt111,Rt112	Pt121,Pt122 Pw121,Pw122 Sr121,Sr122 Rt121,Rt122	Pt131,Pt132 Pw131,Pw132 Sr131,Sr132 Rt131,Rt132

【図9】

入力テキスト	変換後テキスト	自然音声韻律 パターンコード	規則合成韻律目標値 (接続部基本周波数・ 接続音韻最大パワー・ 接続部話速)
「した。」	「しました。」	0000	基本周波数 150Hz-180Hz パワー 60dB 話速 2モーラ/秒
「という。」	「ということです。」	0001	基本周波数 170Hz-190Hz パワー 55dB 話速 1.5モーラ/秒

【図7】



【図8】

